

# Optimal Input Design for Affine Model Discrimination with Applications in Intention-Aware Vehicles

Emil Jacobsen<sup>a,\*</sup>, Farshad Harirchi<sup>b,\*</sup>, Sze Zheng Yong<sup>c,\*</sup> and Necmiye Ozay<sup>b</sup>

**Abstract**—This paper considers the optimal design of input signals for the purpose of discriminating among a finite number of affine models with both controlled and uncontrolled inputs. Each affine model represents a different system operating mode, corresponding to unobserved intents of other drivers or robots, or to fault types or attack strategies, etc. The input design problem aims to find optimal separating/discriminating (controlled) inputs such that the outputs of all the affine models are guaranteed to differ in at least one time instance. We formulate this problem as a tractable mixed-integer linear program (MILP), with an emphasis on guarantees for model discrimination and with no relaxations. Moreover, our fairly general formulation allows the incorporation of differing objectives or responsibilities among rational agents. For instance, each driver has to individually obey traffic rules, while simultaneously optimizing for safety, comfort and energy efficiency. Finally, we demonstrate the effectiveness of our approach for identifying the intent of other human-driven or autonomous vehicles.

## I. INTRODUCTION

As cyber-physical systems become increasingly complex, integrated and interconnected, they inevitably have to interact with other systems under partial knowledge of each other's internal state such as intention and mode of operation. For instance, autonomous vehicles and robots must operate without access to the intentions or decisions of nearby vehicles or humans [1], [2]. Similarly, system behaviors change in the presence of different fault types [3], [4] or attack modes [5], [6], and the true system model is not directly observed. In both examples, there is a number of possible system behaviors and the objective is to develop methods for discriminating between these models (of system behaviors) based on observed measurements. This is an important problem in statistics, machine learning and systems theory; thus, general techniques for model discrimination can have a significant impact on a broad range of applications.

1) *Literature Review*: The problem of discriminating among a set of linear models appears in a wide variety of research areas such as fault detection, input-distinguishability and mode discernibility of hybrid systems. Two categories of approaches have been utilized in the literature: passive and active. While passive methods seek the separation of the

models regardless of the input, active methods search for an input such that the behaviors of different models are distinct. The concept of input-distinguishability for a pair of linear models is discussed in [7], [8], which is the passive version of distinguishability that is defined in [9]. The problem of model-based active fault detection is also extensively studied, where the goal is to find a small excitation that has a minimal effect on the desired behavior of the system, while guaranteeing the isolation of different fault models [10], [4], [11]. On the other hand, a computational method for passively discriminating among fault models is given in [12], [13]. The problem of mode discernibility in switched autonomous linear models in the passive setting is also studied in [14], and in addition, they introduced the concept of controlled-discernibility, which seeks for an input to discriminate among different mode sequences in switched linear models.

2) *Main Contributions*: We propose a novel optimization-based method for *active* model discrimination through the design of optimal separating/discriminating inputs. First, we extend the class of models considered in [10], [11] to a more general class of affine models with unknown inputs. Moreover, we consider a principled characterization of input and state constraints, with a clear description of individual responsibilities for constraint satisfaction when multiple rational agents are involved. To the best extent of our knowledge, our framework is the first to consider the notion of individual responsibilities for constraint satisfaction in active model discrimination.

We formulate the active model discrimination problem as a tractable MILP problem (MIQP, in the case of quadratic objective function). This approach guarantees the separation among affine models. The convex hull reformulation utilized in our previous work [3], [12] to derive an equivalent MILP formulation from a nonlinear mixed integer program is not directly applicable for this problem. Instead, we employ Special Ordered Set of degree 1 (SOS-1) constraints [15] (also considered in our recent work [13]) to express the model discrimination problem as an MILP. In addition, we leverage recent literature on robust optimization [16], [17] to handle model uncertainties, in contrast to existing approaches that consider polyhedral projections [10] and zonotopes [11] that can be rather limiting. Finally, the effectiveness of our approach is illustrated with an application problem of identifying the intention of other human-driven or autonomous vehicles approaching an intersection.

3) *Notation and Definitions*: Let  $\mathbf{x} \in \mathbb{R}^n$  denote a vector and  $\mathbf{M} \in \mathbb{R}^{n \times m}$  a matrix, with transpose  $\mathbf{M}^T$  and  $\mathbf{M} \geq 0$  denotes element-wise non-negativity. The vector norm of  $\mathbf{x}$  is

<sup>a</sup> Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden. (email: emiljaco@kth.se)

<sup>b</sup> Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI, 48109. (email: {harirchi,necmiye}@umich.edu)

<sup>c</sup> School for Engineering in Matter, Transport and Energy, Arizona State University, Tempe, AZ, 85281. (email: sze.zheng.yong@asu.edu)

\* Equal contribution from these authors.

This work was supported by an Early Career Faculty grant from NASA's Space Technology Research Grants Program and DARPA grant N66001-14-1-4045. This work was done at the University of Michigan, Ann Arbor.

denoted by  $\|\mathbf{x}\|_i$  with  $i \in \{1, 2, \infty\}$ , while  $\mathbf{1}$  and  $I$  represent the vector of ones and the identity matrix of appropriate dimensions. The diag and vec operators are defined for a collection of matrices  $\mathbf{M}_i, i = 1, \dots, n$  and matrix  $\mathbf{M}$  as:

$$\text{diag}_{i=1}^n \{\mathbf{M}_i\} = \begin{bmatrix} \mathbf{M}_1 & & \\ & \ddots & \\ & & \mathbf{M}_n \end{bmatrix}, \quad \text{vec}_{i=1}^n \{\mathbf{M}_i\} = \begin{bmatrix} \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_n \end{bmatrix}, \quad (1)$$

$\text{diag}_N \{\mathbf{M}\} = I_N \otimes \mathbf{M}$ ,  $\text{vec}_N \{\mathbf{M}\} = \mathbf{1}_N \otimes \mathbf{M}$ , where  $\otimes$  is the Kronecker product. The set of positive integers up to  $n$  is denoted by  $\mathbb{Z}_n^+$ , and the set of non-negative integers up to  $n$  is denoted by  $\mathbb{Z}_n^0$ . We will also make frequent use of Special Ordered Set of degree 1 (SOS-1) constraints<sup>1</sup> in our optimization formulation, defined as follows:

**Definition 1** (SOS-1 Constraints [18]). *A special ordered set of degree 1 (SOS-1) constraint is a set of integer, continuous or mixed-integer scalar variables for which at most one variable in the set may take a value other than zero, denoted as SOS-1:  $\{v_1, \dots, v_N\}$ . For instance, if  $v_i \neq 0$ , then this constraint imposes that  $v_j = 0$  for all  $j \neq i$ .*

## II. PRELIMINARIES

In this section, we describe the modeling framework we consider, followed by a formal problem statement.

### A. Modeling Framework

Consider  $N$  discrete-time affine time-invariant models  $\mathcal{G}_i = (A_i, B_i, C_i, f_i)$ , each with states,  $\mathbf{x}_i \in \mathbb{R}^n$ , outputs,  $z_i \in \mathbb{R}^p$ , and inputs,  $\mathbf{u}_i \in \mathbb{R}^m$ . The state and output equations of the models are:

$$\mathbf{x}_i(k+1) = A_i \mathbf{x}_i(k) + B_i \mathbf{u}_i(k) + f_i, \quad (2)$$

$$z_i(k) = C_i \mathbf{x}_i(k) + g_i. \quad (3)$$

The initial condition is shared by all the models and is denoted by  $\mathbf{x}_0 = \mathbf{x}_i(0)$ .

The first  $m_u$  components of  $\mathbf{u}_i$  are controlled inputs, denoted as  $u \in \mathbb{R}^{m_u}$ , which are equal for all  $\mathbf{u}_i$ , while the other  $m_w = m - m_u$  components of  $\mathbf{u}_i$ , denoted as  $w_i \in \mathbb{R}^{m-w}$ , are unknown inputs that are model-dependent. Further, the states  $\mathbf{x}_i$  are partitioned into  $x_i \in \mathbb{R}^{n_x}$  and  $y_i \in \mathbb{R}^{n_y}$ , where  $n_y = n - n_x$ , as follows:

$$\mathbf{u}_i(k) = \begin{bmatrix} u(k) \\ w_i(k) \end{bmatrix}, \quad \mathbf{x}_i(k) = \begin{bmatrix} x_i(k) \\ y_i(k) \end{bmatrix}. \quad (4)$$

The states  $x_i$  and  $y_i$  represent the subset of the states  $\mathbf{x}_i$  that are the ‘responsibilities’ of the inputs  $u$  and  $w_i$ , respectively. The term ‘responsibility’ in this paper is to be interpreted as  $u$  and  $w_i$ , respectively, having to independently satisfy the following polyhedral state constraints:

$$x_i(k) \in \mathcal{X}_{x,i} = \{x \in \mathbb{R}^{n_x} : P_{x,i}x \leq p_{x,i}\}, \quad (5)$$

$$y_i(k) \in \mathcal{X}_{y,i} = \{y \in \mathbb{R}^{n_y} : P_{y,i}y \leq p_{y,i}\}. \quad (6)$$

Moreover, the inputs and initial conditions are constrained to polyhedral sets as follows:

$$u(k) \in \mathcal{U} = \{u \in \mathbb{R}^{m_u} : Q_u u \leq q_u\} \quad (7)$$

$$w_i(k) \in \mathcal{W}_i = \{w \in \mathbb{R}^{m_w} : Q_{w,i} w \leq q_{w,i}\} \quad (8)$$

$$\mathbf{x}_0 \in \mathcal{X}_0 = \{x \in \mathbb{R}^n : P_0 x \leq p_0\}. \quad (9)$$

<sup>1</sup>Off-the-shelf solvers such as Gurobi and CPLEX [18], [19] can readily handle these constraints, which can significantly reduce the search space for integer variables in branch and bound algorithms.

It is also noteworthy that the input  $u$  that we will design has to satisfy the state constraints in (5) for all models  $i \in \mathbb{Z}_N^+$ , similar in spirit to robust control problems. On the other hand, the unknown input  $w_i$  has to only satisfy its corresponding state constraints in (6).

Using the above partitions of states and inputs, the corresponding partitioning of the state equations in (2) is:

$$\mathbf{x}_i(k+1) = \begin{bmatrix} A_{xx,i} & A_{xy,i} \\ A_{yx,i} & A_{yy,i} \end{bmatrix} \mathbf{x}_i(k) + \begin{bmatrix} B_{xu,i} & B_{xw,i} \\ B_{yu,i} & B_{yw,i} \end{bmatrix} \mathbf{u}_i(k) + \begin{bmatrix} f_{x,i} \\ f_{y,i} \end{bmatrix}. \quad (10)$$

Further, we will consider a time horizon of length  $T$  and introduce some time-concatenated notation. The time-concatenated states and outputs are defined as

$$\mathbf{x}_{i,T} = \text{vec}_{j=1}^T \{\mathbf{x}_i(j)\}, \quad x_{i,T} = \text{vec}_{j=1}^T \{x_i(j)\}, \\ y_{i,T} = \text{vec}_{j=1}^T \{y_i(j)\}, \quad z_{i,T} = \text{vec}_{j=1}^T \{z_i(j)\},$$

while the time-concatenated inputs are defined as

$$\mathbf{u}_{i,T} = \text{vec}_{j=0}^{T-1} \{\mathbf{u}_i(j)\}, \quad u_T = \text{vec}_{j=0}^{T-1} \{u(j)\}, \quad w_{i,T} = \text{vec}_{j=0}^{T-1} \{w_i(j)\}.$$

Then, concatenating  $\mathbf{x}_{i,T}$ ,  $x_{i,T}$ ,  $y_{i,T}$ ,  $w_{i,T}$  and  $z_{i,T}$  across all modes as

$$\mathbf{x}_T = \text{vec}_{i=1}^N \{\mathbf{x}_{i,T}\}, \quad x_T = \text{vec}_{i=1}^N \{x_{i,T}\}, \quad y_T = \text{vec}_{i=1}^N \{y_{i,T}\}, \\ w_T = \text{vec}_{i=1}^N \{w_{i,T}\}, \quad z_T = \text{vec}_{i=1}^N \{z_{i,T}\}, \quad (11)$$

the states and outputs over the entire time horizon can be written as a simple function of the initial state  $\mathbf{x}_0$  and input vectors  $u_T$  and  $w_T$ :

$$x_T = M_x \mathbf{x}_0 + \Gamma_{xu} u_T + \Gamma_{xw} w_T + \tilde{f}_x, \quad (12)$$

$$y_T = M_y \mathbf{x}_0 + \Gamma_{yu} u_T + \Gamma_{yw} w_T + \tilde{f}_y, \quad (13)$$

$$x_T = \bar{A} \mathbf{x}_0 + \Gamma_u u_T + \Gamma_w w_T + \tilde{f}, \quad (14)$$

$$z_T = \bar{C} x_T + \tilde{g}. \quad (15)$$

The matrices and vectors  $M_\star$ ,  $\Gamma_{\star u}$ ,  $\Gamma_{\star w}$  and  $\tilde{f}_\star$  for  $\star \in \{x, y\}$ , and  $\bar{A}$ ,  $\Gamma_u$ ,  $\Gamma_w$ ,  $\bar{C}$ ,  $\tilde{f}$ ,  $\tilde{g}$  are defined in the appendix.

**Remark 1.** *Since it is the responsibility of  $w_i$  to satisfy the constraint in (6), it is important to make sure that the models are meaningful in the sense that for the range of time horizons of interest,  $T$ , and for each  $i \in \mathbb{Z}_N^+$ ,*

$$\exists w_i(k) \in \mathcal{W}_i, \forall k \in \mathbb{Z}_{T-1}^0 : (6) \text{ is satisfied} \quad (16)$$

*for any given  $\mathbf{x}_0 \in \mathcal{X}_0$  that satisfies  $y_i(0) \in \mathcal{X}_{y,i}$  (cf. (6)) for all  $i \in \mathbb{Z}_N^+$ , and for any given  $u(k) \in \mathcal{U}$  for all  $k \in \mathbb{Z}_{T-1}^0$ . If the considered affine model satisfies this assumption, we refer to it as a well-posed affine model. Note that models that do not satisfy this assumption are impractical, since the responsibilities of the uncontrolled input will be impossible to be satisfied; thus, we shall assume throughout the paper that the given affine models are always well-posed.*

### B. Problem Statement

In order to formally state the problem, let us first define the consistency set for each well-posed affine model.

**Definition 2.** (Consistency Set) *The consistency set for a well-posed affine model  $\mathcal{G}_i$  is defined as follows:*

$$\mathcal{T}_{\mathcal{G}_i}(u_T, w_{i,T}, z_T, \mathbf{x}^*) = \{\mathbf{x} \mid (2), (3) \text{ are satisfied}, \mathbf{x}(0) = \mathbf{x}^*\}. \quad (17)$$

The active model discrimination problem is defined as:

**Problem 1** (Active Model Discrimination). *Find an optimal input sequence  $u_T^*$  subject to  $u(k) \in \mathcal{U}$  that minimizes a predefined cost function  $c(u_T)$  such that for any observed output sequence  $z_T$ ,*

$$\bigcap_{i=1}^N \mathcal{T}_{G_i}(u_T, w_{i,T}, z_T, \mathbf{x}^*) = \emptyset \text{ and (5) hold} \quad (18)$$

$$\forall(\mathbf{x}^*, w_T) \in \{\mathbf{x}^* \in \mathcal{X}_0, w_i(k) \in \mathcal{W}_i : (6) \text{ holds}\},$$

where  $\mathcal{X}_0$ ,  $\mathcal{U}$ ,  $\mathcal{W}_i$  and  $\mathcal{T}_{G_i}$  are given in (9), (7), (8) and (17), respectively.

### III. ACTIVE MODEL DISCRIMINATION APPROACH

In this section, we propose an optimization-based approach to solve Problem 1. First, we will translate the problem described in terms of consistency sets into an equivalent robust optimization problem, which can be further transformed into an MILP for which off-the-shelf optimization softwares are readily available [18], [19]. For the sake of clarity, we will defer the detailed derivations and definitions of matrices in the following section to the appendix.

#### A. Equivalent Robust Optimization Problem

The active model discrimination problem in Problem 1 can be reformulated as a robust optimization problem.

**Proposition 1** (Robust Optimization Formulation). *Given a separability index  $\epsilon$ , the active model discrimination problem in Problem 1 is equivalent up to the separability index  $\epsilon$  to the following:*

$$\min_{u_T, w_T, \mathbf{x}_0, \mathbf{y}_T, \mathbf{z}_T, \mathbf{x}_T, \mathbf{s}, \mathbf{a}} c(u_T) \quad (P_{Robust})$$

$$\text{s.t. } \forall k \in \mathbb{Z}_T^0 : (3)-(5), (7), \forall k \in \mathbb{Z}_{T-1}^0 : (10) \text{ hold,} \quad (19a)$$

$$\forall i, j \in \mathbb{Z}_N^+, i < j, \quad z_{i,l}(k) - z_{j,l}(k) - \epsilon + s_{i,j,k,l,1} \geq 0,$$

$$\forall l \in \mathbb{Z}_p^1, k \in \mathbb{Z}_T^0 : z_{j,l}(k) - z_{i,l}(k) - \epsilon + s_{i,j,k,l,2} \geq 0, \quad (19b)$$

$$\forall \alpha \in \{1, 2\} \quad a_{i,j,k,l,\alpha} \in \{0, 1\},$$

$$\text{SOS-1: } \{s_{i,j,k,l,\alpha}, a_{i,j,k,l,\alpha}\},$$

$$\sum_{k \in \mathbb{Z}_T^0} \sum_{l \in \mathbb{Z}_p^1} \sum_{\alpha \in \{1, 2\}} a_{i,j,k,l,\alpha} \geq 1, \quad (19c)$$

$$\forall k \in \mathbb{Z}_T^0, w_T, \mathbf{x}_0 : (6), (8), (9) \text{ holds,} \quad (19d)$$

where  $\mathbf{a}$  is the vector of binary variables  $a_{i,j,k,l,\alpha}$  concatenated over the indices in the order  $i, j, k, l, \alpha$ , and  $\mathbf{s}$  is similarly a vector of slack variables  $s_{i,j,k,l,\alpha}$ , defined as  $\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$ , where  $s_\alpha$  for  $\alpha \in \{1, 2\}$  are defined in (22) in the appendix.

*Proof.* From Definition 2 and Problem  $P_{Robust}$ , it is straightforward to see that an input sequence  $u_T$  can discriminate among models if and only if the output trajectories of each pair of models differ in at least one time instance for all possible initial states  $\mathbf{x}_0$  and unknown input sequences  $w_T$ , i.e.,  $\forall i, j \in \mathbb{Z}_N^+, i < j, \exists k \in \mathbb{Z}_T^0$  such that  $z_i(k) \neq z_j(k)$  (separability condition). Note that the non-convex separability condition is replaced by a separation of at least  $\epsilon$  (represented by (19b) and (19c)), where  $\epsilon$  is the amount of desired separation or simply the machine precision. To enforce this condition as constraints that are more amenable for use in optimization softwares, condition (19c) is added to enforce that at least one component of  $\mathbf{a}$  is 1. Then, by the SOS-1 constraint in (19b), the corresponding component of  $\mathbf{s}$

has to be 0 by Definition 1. In this case (when  $s_{i,j,k,l,\alpha} = 0$ ), the first two constraints in (19b) impose that  $z_i(k)$  and  $z_j(k)$  differ by at least  $\epsilon$  in the  $l$ -th component.  $\square$

#### B. Equivalent MILP via Robustification

The formulation in Proposition 1 is still not readily implementable, as we still have semi-infinite constraints, in the form of (19d). Thus, we leverage recent literature on robust optimization [16], [17] to convert this problem into a mixed-integer linear program (MILP), which can readily be solved using off-the-shelf solvers such as Gurobi and CPLEX [18], [19]. The main result of the paper is stated as follows:

**Theorem 1** (Discriminating Input Design as an MILP). *Given well-posed affine models and the separability index  $\epsilon$ , consider the following problem:*

$$\min_{u_T, s, \mathbf{a}, \Pi} c(u_T) \quad (P_{DID})$$

$$\text{s.t. } \bar{Q}_u u_T \leq \bar{q}_u, \Phi^T \Pi = R^T, \Pi \geq 0, \quad (20a)$$

$$\Pi^T \phi \leq r(u_T, s), \mathbf{a} \in \{0, 1\}^{pTN(N-1)}, \quad (20b)$$

$$\sum_{k \in \mathbb{Z}_T^0} \sum_{l \in \mathbb{Z}_p^1} \sum_{\alpha \in \{1, 2\}} a_{i,j,k,l,\alpha} \geq 1, \quad (20c)$$

$$\text{SOS-1: } \{s_{i,j,k,l,\alpha}, a_{i,j,k,l,\alpha}\}, \quad (20d)$$

$\forall i, j \in \mathbb{Z}_N^+ : i < j, \forall l \in \mathbb{Z}_p^1, \forall k \in \mathbb{Z}_T^0$  and  $\alpha \in \{1, 2\}$ , where  $\Pi$  is a matrix of dual variables, while  $\bar{Q}_u, \Phi, R, \bar{q}_u, \phi$  and  $r(u_T, s)$  are problem-dependent matrices and vectors that are defined in the appendix. Then,

- 1) if  $\bar{P}_y \Gamma_{yu} = 0$ , Problem  $(P_{DID})$  is equivalent up to the separability index  $\epsilon$  to Problem 1 and its solution is optimal;
- 2) if  $\bar{P}_y \Gamma_{yu} \neq 0$  and if Problem  $(P_{DID})$  is feasible, its solution is sub-optimal,

where  $\bar{P}_y = \text{diag}_{i=1}^N \text{diag}_T\{P_{y,i}\}$  and  $P_{y,i}$  is given in (6).

*Proof.* Provided in the appendix.  $\square$

**Remark 2.** The matrix  $\Gamma_{yu}$  in (13) plays an important role in determining the optimality of the designed input when using our proposed approach. In particular, our approach may only yield a suboptimal input if  $\Gamma_{yu} \neq 0$ , as illustrated by Theorem 1. An effort to reduce the optimality gap is an ongoing subject of our research (see [20]).

### IV. ILLUSTRATIVE EXAMPLE: ACTIVE INTENTION IDENTIFICATION OF OTHER VEHICLES

In this section, we demonstrate the efficacy of the proposed model discrimination approach by applying it to the problem of intention identification of human-driven or autonomous vehicles using a suitable choice of inputs. Our formulation is general enough to capture various driving scenarios such as merging lanes, road intersections, lane changes, etc., provided that suitable models of intentions are available (from first principles or via data-driven approaches). However, for the sake of brevity, we illustrate the proposed approach in the context of active intention identification using only the scenario of multiple vehicles crossing an intersection. The controlled/ego vehicle can choose any control input to identify the intention of other vehicles subject to engine power and braking limitations, as long as traffic laws are

obeyed. On the other hand, the driver of the other vehicles—autonomous or human-driven—can choose among three intentions<sup>2</sup> described in Section IV-A.

The objective is to identify the intention of other *rational* drivers at intersections by taking appropriate control actions while optimizing for safety, comfort or energy efficiency. For the implementation of proposed approach, we utilized Yalmip [21] and Gurobi [18] in the MATLAB environment.

#### A. Description of Vehicle Dynamics and Intention Models

We consider two<sup>3</sup> vehicles at an intersection (origin of coordinate system) with feedback linearized vehicle dynamics:  $\ddot{x} = u$ ,  $\ddot{y} = w$ , with  $x$ ,  $y$  and  $\dot{x}$ ,  $\dot{y}$  being vehicle positions and velocities in  $m$  and  $\frac{m}{s}$ , respectively, while  $u$  and  $w$  are the control inputs of the ego and other vehicles in  $\frac{N}{kg}$ , respectively. The maximum control input of both vehicles (i.e.,  $u$  and  $w$ ) is  $3.97 \frac{N}{kg}$  (corresponding on an acceleration of  $0-100 \frac{km}{h}$  in  $7s$ ), while the minimum input of  $-7.85 \frac{N}{kg}$  corresponds to maximum braking force of  $0.8g$ .

We consider three driver intentions,  $i \in \{I, C, M\}$ , corresponding to Inattentive, Cautious and Malicious drivers. Using  $\mathbf{x} = [x, \dot{x}, y, \dot{y}]$  as the state vector and a sampling time of  $\delta t = 0.3s$ , the vehicle dynamics and intentions are modeled with the following set of parameters and inputs:

**Inattentive Driver** ( $i = I$ ), who fails to notice the ego vehicle, thus proceeding with an unknown time-varying input  $w$  (uncorrelated with  $x$  and  $\dot{x}$ , otherwise unrestricted) that maintains the velocity within a desired range to ensure forward mobility:

$$A_I = \begin{bmatrix} 1 & \delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \delta t \\ 0 & 0 & 0 & 1 \end{bmatrix}, B_I = \begin{bmatrix} 0 & 0 \\ \delta t & 0 \\ 0 & 0 \\ 0 & \delta t \end{bmatrix}, C_I = I_2.$$

The inattentive vehicle maintains forward mobility by ensuring that the velocity satisfies:  $\dot{y}(t) \in [5.56, 13.89] \frac{m}{s}$ .

**Cautious Driver** ( $i = C$ ), who stops at intersection with  $w = -K_p y - K_d \dot{y}$ , where  $K_p = 1.5$  and  $K_d = 4.75$ :

$$A_C = \begin{bmatrix} 1 & \delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \delta t \\ 0 & 0 & -K_p \delta t & 1 - K_d \delta t \end{bmatrix}, B_C = \begin{bmatrix} 0 & 0 \\ \delta t & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, C_C = I_2.$$

**Malicious Driver** ( $i = M$ ), who attempts to cause a collision with  $w = K_p(x - y) + K_d(\dot{x} - \dot{y})$ , where  $K_p = 1$  and  $K_d = 3.5$ :

$$A_M = \begin{bmatrix} 1 & \delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \delta t \\ K_p \delta t & K_d \delta t & -K_p \delta t & 1 - K_d \delta t \end{bmatrix}, B_M = B_C, C_M = I_2.$$

Moreover, the velocity of the ego vehicle is constrained to be between 0 and  $13.89 \frac{m}{s}$  to prevent it from moving backwards. To ensure that the ego vehicle can always stop before the intersection in the event that the intention of the other vehicle cannot be uniquely determined, the ego vehicle is also constrained to never exceed a minimum stopping

distance from the intersection of  $12.29m$ . Notice that with these models of vehicle dynamics and intentions, we have  $\bar{P}_y \Gamma_{yu} = 0$ . This means that the solutions to our active intention identification problem are optimal by Theorem 1.

The main objective is to find an intention-revealing input for differentiating among potential driver intentions *offline*, since online approaches may be computationally prohibitive. The rationale is that even if better inputs can be computed online, we can always fall back on the offline solution when the online approach does not find a solution in a timely manner. However, since we will not know the initial states (i.e., position and velocities) of both vehicles *a priori*, we need to solve a harder problem of designing an optimal input that guarantees intention identification for all possible initial states. In this example, we choose the set of initial states as:

$$\begin{aligned} x(0) &\in [-13.27, -8.85] m, & \dot{x}(0) &\in [20, 30] km/h, \\ y(0) &\in [-100, 0] m, & \dot{y}(0) &\in [20, 50] km/h. \end{aligned} \quad (21)$$

#### B. Active Intention Identification Problem

Using the above models of vehicle dynamics and intentions, we can pose the active intention identification problem as an active model discrimination problem. Specifically, the *offline* active intention identification problem is equivalent to finding  $u_T$  such that  $\bigcap_{i \in \{I, C, M\}} \mathcal{T}_{G_i}(u_T, w_{i,T}, z_T, \mathbf{x}_0) = \emptyset$  for all initial states  $\mathbf{x}_0$  that satisfy (21) (cf. Problem 1). Note that our model discrimination approach can also be applied to *online* intention identification, since this is equivalent to finding inputs  $u_T$  such that  $\bigcap_{i \in \{I, C, M\}} \mathcal{T}_{G_i}(u_T, w_{i,T}, z_T, \mathbf{x}_0) = \emptyset$  with the current state as  $\mathbf{x}_0$ .

Fig. 1 shows an illustration of a potential scenario that facilitates intention identification. In this case, the ego vehicle slows down enough in order to differentiate between the inattentive and malicious drivers but does not come to a stop like the cautious driver. It is noteworthy that the ego vehicle has to apply a non-zero control input because otherwise, there will be no way to distinguish between an inattentive and a malicious driver, since the inattentive driver is free to choose any control inputs.

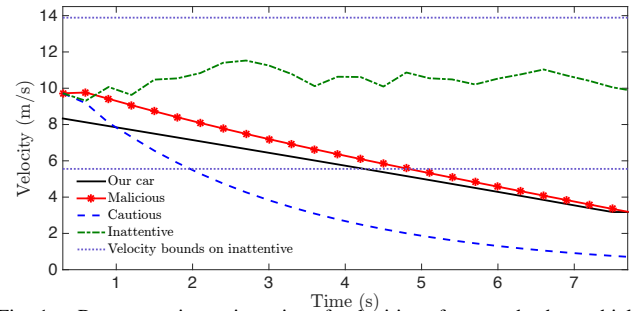


Fig. 1. Representative trajectories of velocities of ego and other vehicles under various intentions.

#### C. A Comparison of Various Objective Functions

In addition to designing an intention-revealing input, our model discrimination approach allows us to choose any convex objective functions to represent driver preferences such as comfort and fuel efficiency. Various objective functions are considered for the input sequence in this case study. For instance,  $\|u_T\|_1$  enforces sparsity in the solution

<sup>2</sup>The assumed intentions (also used in [2]) are for illustrative purposes only; the proposed approach is also applicable with other intention models.

<sup>3</sup>This is for ease of exposition. In fact, *arbitrary* number of vehicles can be handled with the same input as long as their initial conditions are within a predefined set,  $\mathcal{X}_0$  (e.g., within a certain distance before the intersection).

(corresponding to minimal number of non-zero inputs), while  $\|u_T\|_2$  minimizes fuel consumption and  $\|u_T\|_\infty$  ensures comfort (corresponding to minimal input amplitudes). We also consider combinations of  $\|u_T\|_1$  and  $\|(\Delta u)_T\|_\infty$ , that trade-off between sparsity and comfort, where  $(\Delta u)_T$  denotes the rate of change in the inputs, defined as  $(\Delta u)_T = [u(1) - u(0) \dots u(T-1) - u(T-2)]^T$ .

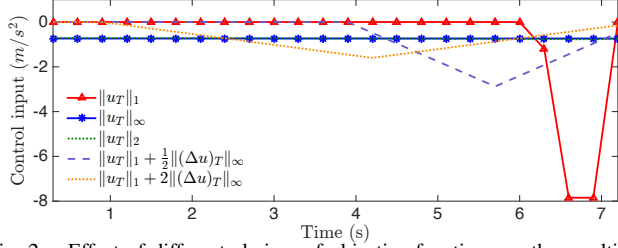


Fig. 2. Effect of different choices of objective functions on the resulting optimal inputs. (Note that the control inputs with objective functions  $\|u_T\|_2$  and  $\|u_T\|_\infty$  are near identical.)

For illustration purposes, we investigated the effect of the objective function choice on the resulting optimal input sequence. As seen in Fig. 2, the optimal input that minimizes  $\|u_T\|_1$  brakes hard at a few time instances while choosing zero inputs at other times. On the other hand, the optimal inputs that minimize  $\|u_T\|_2$  and  $\|u_T\|_\infty$  involve little braking, but at all times, while mixtures of  $\|u_T\|_1$  and  $\|(\Delta u)_T\|_\infty$  provide intermediate levels of braking over some subsets of the time horizon. Note that if a desired control input sequence is given, our approach can also minimize deviations from this input sequence while guaranteeing intention identification.

## V. CONCLUSION

In this paper, a model-based optimization approach is proposed to find a safe and optimal discriminating input, which guarantees the distinction amongst multiple affine models with unknown inputs. The considered modeling framework is fairly general, as it allows the unknown inputs to be applied by interactive rational agents who are responsible for obeying some rules represented by polyhedral constraints. We posed this problem as a tractable mixed-integer linear program (MILP), which can be solved using off-the-shelf optimization softwares. Furthermore, our model discrimination formulation is, to the extent of our knowledge, the first that takes into account different responsibilities among rational agents. To illustrate the efficacy of this framework, we successfully applied the proposed active model discrimination approach to the problem of intention identification of other human-driven or autonomous vehicles at intersections. For future work, we are interested to overcome the limitation of our approach as discussed in Remark 2 to lower the optimality gap of our solutions, while providing guarantees for model discrimination. We will also consider modifications to our approach for efficient real-time implementation.

## APPENDIX: PROOF OF THEOREM 1

To prove Theorem 1, we shall show that Problem  $(P_{DID})$  is equivalent to the robust optimization formulation in Propo-

sition 1. However, before we can proceed, all the matrices required to obtain (12)-(15) are presented below:

$$B_{u,i} = \begin{bmatrix} B_{xu,i} \\ B_{yu,i} \end{bmatrix}, B_{w,i} = \begin{bmatrix} B_{xw,i} \\ B_{yw,i} \end{bmatrix}, \Theta_{i,T} = \begin{bmatrix} I & 0 & \dots & 0 \\ A_i & I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_i^{T-1} & A_i^{T-2} & \dots & I \end{bmatrix},$$

$$\tilde{f}_{\dagger,i,T} = \text{vec}_{i=1}^N \{\tilde{f}_{\dagger,i,T}\}, \dagger \in \{x, y\}$$

$$\bar{A}_{i,T} = \begin{bmatrix} A_i \\ A_i^2 \\ \vdots \\ A_i^T \end{bmatrix}, \Gamma_{\star,i,T} = \begin{bmatrix} B_{\star,i} & & & \\ A_i B_{\star,i} & & & \\ \vdots & & & \\ A_i^{T-1} B_{\star,i} & A_i^{T-2} B_{\star,i} & \dots & B_{\star,i} \end{bmatrix},$$

$$\star \in \{u, w\}, \bar{A} = \text{vec}_{i=1}^N \{\bar{A}_{i,T}\}, \Gamma_u = \text{vec}_{i=1}^N \{\Gamma_{u,i,T}\},$$

$$\tilde{f} = \text{vec}_{i=1}^N \{\tilde{f}_{i,T}\}, \Gamma_w = \text{diag}_{i=1}^N \{\Gamma_{w,i,T}\},$$

$$A_{x,d,i,T} = \text{diag}_T \{A_{xx,i} \ A_{xy,i}\}, B_{xu,d,i,T} = \text{diag}_T \{B_{xu,i}\},$$

$$B_{x,d,i,T} = \text{diag}_T \{B_{xu,i} \ B_{xw,i}\}, B_{xw,d,i,T} = \text{diag}_T \{B_{xw,i}\},$$

$$M_{x,i,T} = A_{x,d,i,T} \begin{bmatrix} I \\ \bar{A}_{i,T-1} \end{bmatrix}, \bar{x} = \begin{bmatrix} x_0 \\ w_T \end{bmatrix}$$

$$\tilde{f}_{x,i,T} = A_{x,d,i,T} \begin{bmatrix} 0 \\ \Theta_{i,T-1} \end{bmatrix} \tilde{f}_{i,T-1} + \tilde{f}_{x,i,T},$$

$$\Gamma_{xu,i,T} = A_{x,d,i,T} \begin{bmatrix} 0 & 0 \\ \Gamma_{xu,T-1} & 0 \end{bmatrix} + B_{xu,d,i,T},$$

$$\Gamma_{xw,i,T} = A_{x,d,i,T} \begin{bmatrix} 0 & 0 \\ \Gamma_{xw,T-1} & 0 \end{bmatrix} + B_{xw,d,i,T}$$

$$M_x = \text{vec}_{i=1}^N \{M_{x,i,T}\}, \Gamma_{xu} = \text{vec}_{i=1}^N \{\Gamma_{xu,i,T}\},$$

$$\tilde{f}_x = \text{vec}_{i=1}^N \{\tilde{f}_{x,i,T}\}, \Gamma_{xw} = \text{diag}_{i=1}^N \{\Gamma_{xw,i,T}\},$$

$$\bar{C} = \text{diag}_{i=1}^N \text{diag}_T \{C_i\}, \tilde{g}_{i,T} = \text{vec}_T \{g_i\}, \tilde{g} = \text{vec}_{i=1}^N \{\tilde{g}_{i,T}\}.$$

Next, we provide some definitions that are necessary for the proof.

1) *Definitions related to state and input constraints:* We shall start by concatenating the polyhedral state constraints in (5) and (6), eliminating  $x_T$  and  $y_T$  in them and replacing  $x_0$  and  $w_T$  by  $\bar{x}$ . First, let

$$\bar{P}_x = \text{diag}_{i=1}^N \text{diag}_T \{P_{x,i}\}, \bar{P}_y = \text{diag}_{i=1}^N \text{diag}_T \{P_{y,i}\},$$

$$\bar{p}_x = \text{diag}_{i=1}^N \text{diag}_T \{p_{x,i}\}, \bar{p}_y = \text{diag}_{i=1}^N \text{diag}_T \{p_{y,i}\}.$$

Then, we can rewrite the polyhedral constraints on  $\dagger$  as:

$$\bar{P}_{\dagger} x_T \leq \bar{p}_{\dagger} \Leftrightarrow H_{\dagger} \bar{x} \leq h_{\dagger}(u_T), \dagger \in \{x, y\},$$

where  $H_{\dagger} = \bar{P}_{\dagger} [M_{\dagger} \ \Gamma_{\dagger w}]$ ,  $h_{\dagger}(u_T) = \bar{p}_{\dagger} - \bar{P}_{\dagger} \Gamma_{\dagger u} u_T - \bar{P}_{\dagger} \tilde{f}_{\dagger}$ .

Similarly, let

$$\bar{Q}_u = \text{diag}_T \{Q_u\}, \bar{Q}_w = \text{diag}_{i=1}^N \text{diag}_T \{Q_{w,i}\},$$

$$\bar{q}_u = \text{diag}_T \{q_u\}, \bar{q}_w = \text{diag}_{i=1}^N \text{diag}_T \{q_{w,i}\}.$$

Then, the polyhedral input constraints in (7) and (8) for all  $k$  are equivalent to  $\bar{Q}_u u_T \leq \bar{q}_u$  and  $\bar{Q}_w w_T \leq \bar{q}_w$ .

Moreover, we have the initial state constraint in (9). Hence, in terms of  $\bar{x}$ , we have a polyhedral constraint of the form

$$H_{\bar{x}} \bar{x} \leq h_{\bar{x}}, \text{ with } H_{\bar{x}} = \begin{bmatrix} P_0 & 0 \\ 0 & \bar{Q}_w \end{bmatrix}, h_{\bar{x}} = \begin{bmatrix} p_0 \\ \bar{q}_w \end{bmatrix}.$$

2) *Definitions related to the separability condition:* We wish to pose the separability condition in the first two equations of (19b) in terms of  $u_T$ ,  $s$ ,  $\bar{x}$ , instead of explicitly using the  $i, j, k$  or  $l$  indices. To do this, we define

$$g = \begin{bmatrix} \tilde{g}_1 - \tilde{g}_2 \\ \tilde{g}_1 - \tilde{g}_3 \\ \vdots \\ \tilde{g}_{N-1} - \tilde{g}_N \end{bmatrix}, E = \begin{bmatrix} E_1 & -E_2 & 0 & \dots & \dots & 0 \\ E_1 & 0 & -E_3 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & E_{N-1} & -E_N \end{bmatrix},$$

$$s_{i,j,\alpha} = \begin{bmatrix} \text{vec}_{l=1}^p \{s_{i,j,1,l,\alpha}\} \\ \vdots \\ \text{vec}_{l=1}^p \{s_{i,j,T,l,\alpha}\} \end{bmatrix}, \quad s_\alpha = \begin{bmatrix} \text{vec}_{j=2}^N \{s_{1,j,\alpha}\} \\ \text{vec}_{j=3}^N \{s_{2,j,\alpha}\} \\ \vdots \\ s_{N-1,N,\alpha} \end{bmatrix} \quad (22)$$

with  $E_i = \text{diag}_T\{C_i\}$ . Then, the separability conditions can be written as  $E\mathbf{x}_T \geq \epsilon\mathbf{1} - s_1 - g$ ,  $-E\mathbf{x}_T \geq \epsilon\mathbf{1} - s_2 + g$ , or equivalently, as:

$$\begin{bmatrix} E \\ -E \end{bmatrix} \mathbf{x}_T = \bar{E}\mathbf{x}_T \geq \epsilon\mathbf{1} - s - \begin{bmatrix} g \\ -g \end{bmatrix} = \epsilon\mathbf{1} - s - \bar{g},$$

where  $\mathbf{1}$  is a vector of ones, of appropriate length. By defining  $\Lambda = \bar{E} \begin{bmatrix} \bar{A} & \Gamma_w \end{bmatrix}$ ,  $\lambda(u_T, s) = \epsilon\mathbf{1} - s - \bar{g} - \bar{E}\Gamma_u u_T - \bar{E}\bar{f}$ , we can summarize the separability inequalities as:

$$\Lambda\bar{x} \geq \lambda(u_T, s).$$

With the above definitions, let us now concatenate the inequalities associated with  $\bar{x}$ , according to whether they are the responsibility of  $u_T$  or of  $w_T$ , given by

$$R\bar{x} \leq r(u_T, s), \quad [\text{Responsibility of } u_T] \quad (23)$$

$$\Phi\bar{x} \leq \psi(u_T), \quad [\text{Responsibility of } w_T] \quad (24)$$

where we defined

$$R = \begin{bmatrix} -\Lambda \\ H_x \end{bmatrix}, r(u_T, s) = \begin{bmatrix} -\lambda(u_T, s) \\ h_x(u_T) \end{bmatrix}, \Phi = \begin{bmatrix} H_y \\ H_{\bar{x}} \end{bmatrix}, \psi(u_T) = \begin{bmatrix} h_y(u_T) \\ h_{\bar{x}} \end{bmatrix}.$$

Finally, we can succinctly express the robust optimization problem in Proposition 1 as follows:

$$\begin{aligned} \min_{u_T, s, a} \quad & c(u_T) \\ \text{s.t.} \quad & \bar{Q}_u u_T \leq \bar{q}_u, R\bar{x} \leq r(u_T, s), \end{aligned} \quad (25a)$$

$$a \in \{0, 1\}^{pTN(N-1)}, \sum_{k,l,\alpha} a_{i,j,k,l,\alpha} \geq 1, \quad (25b)$$

$$\text{SOS-1: } \{s_{i,j,k,l,\alpha}, a_{i,j,k,l,\alpha}\}, \quad (25c)$$

$$\forall \bar{x} : \Phi\bar{x} \leq \psi(u_T). \quad (25d)$$

*Proof of Theorem 1.* Now that we have a succinct formulation of the robust optimization problem in (25), we can utilize ideas presented in [16], [17], [22] to convert the semi-infinite constraints in (25d) into a tractable formulation that is readily implementable using off-the-shelf optimization software. Specifically, since we have a polyhedral uncertainty set (as in [17, p.7] and [22, Lemma 2]), we have

$$\begin{cases} R\bar{x} \leq r(u_T, s), \\ \forall \bar{x} : \Phi\bar{x} \leq \psi(u_T), \end{cases} \Leftrightarrow \begin{cases} \Pi^T \psi(u_T) \leq r, \\ \Phi^T \Pi = R^T, \\ \Pi \geq 0, \end{cases}$$

where  $\Pi$  is matrix of dual variables. Thus, we obtain the following robust counterpart to the formulation in (25):

$$\begin{aligned} \min_{u_T, s, a, \Pi} \quad & c(u_T) \\ \text{s.t.} \quad & \bar{Q}_u u_T \leq \bar{q}_u, \Phi^T \Pi = R^T, \Pi \geq 0, \end{aligned} \quad (26a)$$

$$\Pi^T \psi(u_T) \leq r(u_T, s), \quad (26b)$$

$$a \in \{0, 1\}^{pTN(N-1)}, \sum_{k,l,\alpha} a_{i,j,k,l,\alpha} \geq 1, \quad (26c)$$

$$\text{SOS-1: } \{s_{i,j,k,l,\alpha}, a_{i,j,k,l,\alpha}\}. \quad (26d)$$

If the robust counterpart in (26) were an MILP, then the proof would be complete. However, note that we have a bilinear term on the left hand side of (26b) since  $\psi(u_T)$  is linear in  $u_T$ . Nonetheless, if  $\bar{P}_y \Gamma_{yu} = 0$ , it can be verified that  $\psi(u_T)$  will be independent of  $u_T$ , and we denote this with  $\psi(u_T) = \phi$ . Hence, the first implication in Theorem 1 holds directly.

On the other hand, if  $\bar{P}_y \Gamma_{yu} \neq 0$ , we will choose to restrict the feasibility set of  $u_T$  by relaxing the feasibility

set of  $\bar{x}$  in (24). The reason for considering this restriction is to guarantee that any eventual solution to the new problem is still a valid separating input, although it would be sub-optimal. Looking at condition (24) before robustification, we can relax this inequality for  $\bar{x}$  by overapproximating  $\psi(u_T)$  with  $\phi$ , which is defined as the component-wise maximum of  $\psi(u_T)$  where  $u_T$  is subject to  $u(k) \in \mathcal{U}$  for all  $k \in \mathbb{Z}_{T-1}^0$ . This component-wise maximum can be easily found using a linear program. Since  $\phi$  is, by construction, independent of  $u_T$ , the robust counterpart to this modified problem is an MILP and can be tractably solve with off-the-shelf optimization solvers. Thus, if the modified problem is feasible, then the solution is, by construction, sub-optimal, and so, the second implication in Theorem 1 holds.  $\square$

## REFERENCES

- [1] D. Sadigh, S.S. Sastry, S. Seshia, and A. Dragan. Information gathering actions over human internal state. In *IEEE/RSJ IROS*, Oct. 2016.
- [2] S.Z. Yong, M. Zhu, and E. Frazzoli. Generalized innovation and inference algorithms for hidden mode switched linear stochastic systems with unknown inputs. In *IEEE CDC*, pages 3388–3394, Dec. 2014.
- [3] F. Harirchi and N. Ozay. Model invalidation for switched affine systems with applications to fault and anomaly detection. *IFAC-PapersOnLine*, 48(27):260–266, 2015.
- [4] S. Cheong and I.R. Manchester. Input design for discrimination between classes of LTI models. *Automatica*, 53:103–110, 2015.
- [5] F. Pasqualetti, F. Dörfler, and F. Bullo. Attack detection and identification in cyber-physical systems. *IEEE Trans. on Aut. Cont.*, 58(11):2715–2729, November 2013.
- [6] S.Z. Yong, M. Zhu, and E. Frazzoli. Resilient state estimation against switching attacks on stochastic cyber-physical systems. In *IEEE CDC*, pages 5162–5169, Dec. 2015.
- [7] H. Lou and P. Si. The distinguishability of linear control systems. *Nonlinear Analysis: Hybrid Systems*, 3(1):21–38, 2009.
- [8] P. Rosa and C. Silvestre. On the distinguishability of discrete linear time-invariant dynamic systems. In *IEEE CDC-ECC*, pages 3356–3361, 2011.
- [9] M. Grewal and K. Glover. Identifiability of linear and nonlinear dynamical systems. *IEEE Trans. Aut. Cont.*, 21(6):833–837, 1976.
- [10] R. Nikoukhah and S. Campbell. Auxiliary signal design for active failure detection in uncertain linear systems with a priori information. *Automatica*, 42(2):219–228, 2006.
- [11] J.K. Scott, R. Findeisen, R. D Braatz, and D.M. Raimondo. Input design for guaranteed fault diagnosis using zonotopes. *Automatica*, 50(6):1580–1589, 2014.
- [12] F. Harirchi and N. Ozay. Guaranteed model-based fault detection in cyber-physical systems: A model invalidation approach. *arXiv:1609.05921 [math.OA]*, 2016.
- [13] F. Harirchi and S.Z. Yong. Guaranteed fault detection and isolation for switched affine models, 2017. preprint.
- [14] M. Babaali and M. Egerstedt. Observability of switched linear systems. In *International Workshop on Hybrid Systems: Computation and Control*, pages 48–63. Springer, 2004.
- [15] E. Beale and J. Forrest. Global optimization using special ordered sets. *Mathematical Programming*, 10(1):52–69, 1976.
- [16] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University Press, 2009.
- [17] D. Bertsimas, D.B. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.
- [18] Inc. Gurobi Optimization. Gurobi optimizer reference manual, 2015.
- [19] IBM ILOG CPLEX. V12. 1: User’s manual for cplex. *International Business Machines Corporation*, 46(53):157, 2009.
- [20] F. Harirchi, S.Z. Yong, E. Jacobsen, and N. Ozay. Active model discrimination with applications to fraud detection in smart buildings, 2017. preprint.
- [21] J. Löfberg. Yalmip : A toolbox for modeling and optimization in MATLAB. In *CACSD*, Taipei, Taiwan, 2004.
- [22] D. Bertsimas and F.J.C.T. de Ruiter. Duality in two-stage adaptive linear optimization: Faster computation and stronger bounds. *INFORMS Journal on Computing*, 28(3):500–511, 2016.